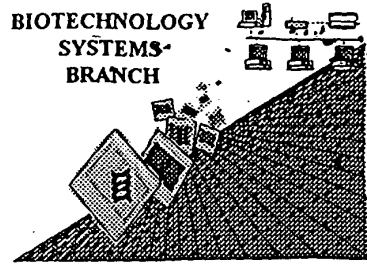


**RAW SEQUENCE LISTING**  
**ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number:

09/835996

Source:

OIPB

Date Processed by STIC:

10/16/01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4211

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER:

09/835 996

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY P1

- 1      Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length.      Sequence(s)          contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)         . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)      Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence  
    (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    (i)      SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
    (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped  
  
    Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)      Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence:  
    <210> sequence id number  
    <400> sequence id number  
    000
- 9      Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents
- 10      Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Artificial Sequence
- 11      Use of <220>      Sequence(s)          missing the <220> "Feature" and associated numeric identifiers and responses.  
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n      n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

## TIME: 15:43:12

Output Set: N:\CRF3\10162001\I835996.raw

Prange on Ref'd 222 does not cover the unknown we presented.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/835,996

DATE: 10/16/2001

TIME: 15:43:12

Input Set : A:\35915a.seq.txt

Output Set: N:\CRF3\10162001\I835996.raw

```

63 5          10          15          20
65 acc cag gca cgg aaa ggc ttc tgg gac tac ttc agc cag acc agc ggg      153
66 Thr Gln Ala Arg Lys Gly Phe Trp Asp Tyr Phe Ser Gln Thr Ser Gly
67          25          30          35
69 gac aaa ggc agg gtg gag cag atc cat cag cag aag atg gct cgc gag      201
70 Asp Lys Gly Arg Val Glu Gln Ile His Gln Gln Lys Met Ala Arg Glu
71          40          45          50
73 ccc gcg acc ctg aaa gac agc ctt gag caa gac ctc aac aat atg aac      249
74 Pro Ala Thr Leu Lys Asp Ser Leu Glu Gln Asp Leu Asn Asn Met Asn
75          55          60          65
77 aag ttc ctg gaa aag ctg agg cct ctg agt ggg agc gag gct cct cgg      297
78 Lys Phe Leu Glu Lys Leu Arg Pro Leu Ser Gly Ser Glu Ala Pro Arg
79          70          75          80
81 ctc cca cag gac ccg gtg ggc atg cgg cgg cag ctg cag gag gag ttg      345
82 Leu Pro Gln Asp Pro Val Gly Met Arg Arg Gln Leu Gln Glu Glu Leu
83 85          90          95          100
85 gag gag gtg aag gct cgc ctc cag ccc tac atg gca gag gcg cac gag      393
86 Glu Glu Val Lys Ala Arg Leu Gln Pro Tyr Met Ala Glu Ala His Glu
87          105          110          115
89 ctg gtg ggc tgg aat ttg gag ggc ttg cgg cag caa ctg aag ccc tac      441
90 Leu Val Gly Trp Asn Leu Glu Gly Leu Arg Gln Gln Leu Lys Pro Tyr
91          120          125          130
93 acg atg gat ctg atg gag cag gtg gcc ctg cgc gtg cag gag ctg cag      489
94 Thr Met Asp Leu Met Glu Gln Val Ala Leu Arg Val Gln Glu Leu Gln
95          135          140          145
97 gag cag ttg cgc gtg gtg ggg gaa gac acc aag gcc cag ttg ctg ggg      537
98 Glu Gln Leu Arg Val Val Gly Glu Asp Thr Lys Ala Gln Leu Leu Gly
99          150          155          160
101 ggc gtg gac gag gct tgg gct ttg ctg cag gga ctg cag agc cgc gtg      585
102 Gly Val Asp Glu Ala Trp Ala Leu Leu Gln Gly Leu Gln Ser Arg Val
103 165          170          175          180
105 gtg cac cac acc ggc cgc ttc aaa gag ctc ttc cac cca tac gcc gag      633
106 Val His His Thr Gly Arg Phe Lys Glu Leu Phe His Pro Tyr Ala Glu
107          185          190          195
109 agc ctg gtg agc ggc atc ggg cgc cac gtg cag gag ctg cac cgc agt      681
110 Ser Leu Val Ser Gly Ile Gly Arg His Val Gln Glu Leu His Arg Ser
111          200          205          210
113 gtg gct ccg cac gcc ccc gcc agc ccc gcg cgc ctc agt cgc tgc gtg      729
114 Val Ala Pro His Ala Pro Ala Ser Pro Ala Arg Leu Ser Arg Cys Val
115          215          220          225
117 cag gtg ctc tcc cgg aag ctc acg ctc aag gcc aag gcc ctg cac gca      777
118 Gln Val Leu Ser Arg Lys Leu Thr Leu Lys Ala Lys Ala Leu His Ala
119          230          235          240
121 cgc atc cag cag aac ctg gac cag ctg cgc gaa gag ctc agc aga gcc      825
122 Arg Ile Gln Gln Asn Leu Asp Gln Leu Arg Glu Glu Leu Ser Arg Ala
123 245          250          255          260
125 ttt gca ggc act ggg act gag gaa ggg gcc ggc ccg gac ccc cag atg      873
126 Phe Ala Gly Thr Gly Thr Glu Glu Gly Ala Gly Pro Asp Pro Gln Met
127          265          270          275

```

## RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/835,996

TIME: 15:43:12

Input Set : A:\35915a.seq.txt

Output Set: N:\CRF3\10162001\I835996.raw

```

129 ctc tcc gag gag gtg cgc cag cga ctt cag gct ttc cgc cag gac acc      921
130 Leu Ser Glu Glu Val Arg Gln Arg Leu Gln Ala Phe Arg Gln Asp Thr
131          280          285          290
133 tac ctg cag ata gct gcc ttc act cgc gcc atc gac cag gag act gag      969
134 Tyr Leu Gln Ile Ala Ala Phe Thr Arg Ala Ile Asp Gln Glu Thr Glu
135          295          300          305
137 gag gtc cag cag cag ctg gcg cca cct cca cca ggc cac agt gcc ttc      1017
138 Glu Val Gln Gln Gln Leu Ala Pro Pro Pro Pro Gly His Ser Ala Phe
139          310          315          320
141 gcc cca gag ttt caa caa aca gac agt ggc aag gtt ctg agc aag ctg      1065
142 Ala Pro Glu Phe Gln Gln Thr Asp Ser Gly Lys Val Leu Ser Lys Leu
143 325          330          335          340
145 cag gcc cgt ctg gat gac ctg tgg gaa gac atc act cac agc ctt cat      1113
146 Gln Ala Arg Leu Asp Asp Leu Trp Glu Asp Ile Thr His Ser Leu His
147          345          350          355
149 gac cag ggc cac agc cat ctg ggg gac ccc tgaggatcta cctgccagg      1163
150 Asp Gln Gly His Ser His Leu Gly Asp Pro
151          360          365
153 cccattccca gctcctgtgc tggggagcct tggctctgag cctctagcat ggttcagtc      1223
155 ttgaaagtgg cctgttgggt ggagggtgga aggtcctgtg caggacaggg aggccaccaa      1283
157 aggggctgct gtctcctgca tatccagcct cctgcgactc cccaatctgg atgcattaca      1343
159 ttcaccaggc tttgcaaacc cagcctccca gtgctcattt gggaatgctc atgagttact      1403
161 ccattcaagg gtgagggagt agggagggag aggcaccatg catgtgggtg attatctgca      1463
163 agcctgtttg ccgtgatgct ggaagcctgt gccactacat cctggagttt ggctctagtc      1523
165 acttctggct gcctgtggc cactgctaca gctggtccac agagaggagc acttgtctcc      1583
167 ccagggtgct catggcagct atcaggggaa tagaaggag aaagagaata tcatggggag      1643
169 aacatgtgat ggtgtgtgaa tatecctgct ggctctgatg ctggtgggta cgaaagggtg      1703
W--> 171 gggctgggat aagagagggc agagcccatg ttttctgaca taactctaca cctatataag      1763*
173 ggactgaacc cttccaactg cgggagctcc ttaaacctt ctggggagca taactggggc      1823
175 tcttcccat cttcagcccc ttctctggg tttcc      1858
178 <210> SEQ ID NO: 2
179 <211> LENGTH: 366
180 <212> TYPE: PRT
181 <213> ORGANISM: Homo sapiens
183 <220> FEATURE:
184 <221> NAME/KEY: misc_feature
185 <222> LOCATION: (46)..(1143)
186 <223> OTHER INFORMATION: n = a or c or g or t.
188 <400> SEQUENCE: 2
190 Met Ala Ser Met Ala Ala Val Leu Thr Trp Ala Leu Ala Leu Leu Ser
191 1          5          10          15
194 Ala Phe Ser Ala Thr Gln Ala Arg Lys Gly Phe Trp Asp Tyr Phe Ser
195          20          25          30
198 Gln Thr Ser Gly Asp Lys Gly Arg Val Glu Gln Ile His Gln Gln Lys
199          35          40          45
202 Met Ala Arg Glu Pro Ala Thr Leu Lys Asp Ser Leu Glu Gln Asp Leu
203          50          55          60
206 Asn Asn Met Asn Lys Phe Leu Glu Lys Leu Arg Pro Leu Ser Gly Ser
207 65          70          75          80

```

\* n at location 1758 is not enumerated in fields 221, 222 and 223.

## RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/835,996

TIME: 15:43:12

Input Set : A:\35915a.seq.txt

Output Set: N:\CRF3\10162001\I835996.raw

```

210 Glu Ala Pro Arg Leu Pro Gln Asp Pro Val Gly Met Arg Arg Gln Leu
211      85      90      95
214 Gln Glu Glu Leu Glu Glu Val Lys Ala Arg Leu Gln Pro Tyr Met Ala
215      100      105      110
218 Glu Ala His Glu Leu Val Gly Trp Asn Leu Glu Gly Leu Arg Gln Gln
219      115      120      125
222 Leu Lys Pro Tyr Thr Met Asp Leu Met Glu Gln Val Ala Leu Arg Val
223      130      135      140
226 Gln Glu Leu Gln Glu Gln Leu Arg Val Val Gly Glu Asp Thr Lys Ala
227 145      150      155      160
230 Gln Leu Leu Gly Gly Val Asp Glu Ala Trp Ala Leu Leu Gln Gly Leu
231      165      170      175
234 Gln Ser Arg Val Val His His Thr Gly Arg Phe Lys Glu Leu Phe His
235      180      185      190
238 Pro Tyr Ala Glu Ser Leu Val Ser Gly Ile Gly Arg His Val Gln Glu
239      195      200      205
242 Leu His Arg Ser Val Ala Pro His Ala Pro Ala Ser Pro Ala Arg Leu
243      210      215      220
246 Ser Arg Cys Val Gln Val Leu Ser Arg Lys Leu Thr Leu Lys Ala Lys
247 225      230      235      240
250 Ala Leu His Ala Arg Ile Gln Gln Asn Leu Asp Gln Leu Arg Glu Glu
251      245      250      255
254 Leu Ser Arg Ala Phe Ala Gly Thr Gly Thr Glu Glu Gly Ala Gly Pro
255      260      265      270
258 Asp Pro Gln Met Leu Ser Glu Glu Val Arg Gln Arg Leu Gln Ala Phe
259      275      280      285
262 Arg Gln Asp Thr Tyr Leu Gln Ile Ala Ala Phe Thr Arg Ala Ile Asp
263      290      295      300
266 Gln Glu Thr Glu Glu Val Gln Gln Gln Leu Ala Pro Pro Pro Pro Gly
267 305      310      315      320
270 His Ser Ala Phe Ala Pro Glu Phe Gln Gln Thr Asp Ser Gly Lys Val
271      325      330      335
274 Leu Ser Lys Leu Gln Ala Arg Leu Asp Asp Leu Trp Glu Asp Ile Thr
275      340      345      350
278 His Ser Leu His Asp Gln Gly His Ser His Leu Gly Asp Pro
279      355      360      365
282 <210> SEQ ID NO: 3
283 <211> LENGTH: 1425
284 <212> TYPE: DNA
285 <213> ORGANISM: Homo sapiens
287 <220> FEATURE:
288 <221> NAME/KEY: CDS
289 <222> LOCATION: (181)..(1146)
291 <400> SEQUENCE: 3
292 gccaggagcc atgtgggttt ttctaggaac caaaatcact tcccggaatt gaccaactgg      60
294 tagactcgcc tagaggggaa gcattgtgtc ctagttagg ctaacagtca gtatccagcc      120
296 tcaacattca gcagaggccc cagatcagcg tctgagccag gccaacaatg accaaggagg      180
298 atg gga tcc tgg gtg cag ctc atc aca agc gtc ggg gtg cag caa aac      228
299 Met Gly Ser Trp Val Gln Leu Ile Thr Ser Val Gly Val Gln Gln Asn

```

## RAW SEQUENCE LISTING

DATE: 10/16/2001

PATENT APPLICATION: US/09/835,996

TIME: 15:43:12

Input Set : A:\35915a.seq.txt

Output Set: N:\CRF3\10162001\I835996.raw

300	1		5		10		15		
302	cat	cca	ggc	tgg	aca	gtg	gct	gga	cag
303	His	Pro	Gly	Trp	Thr	Val	Ala	Gly	Gln
304			20				25		30
306	act	gaa	gaa	gtc	att	gaa	tac	ttc	cag
307	Thr	Glu	Glu	Val	Ile	Glu	Tyr	Phe	Gln
308			35				40		45
310	ctg	aaa	atc	ctg	ctg	act	agc	gat	gaa
311	Leu	Lys	Ile	Leu	Leu	Thr	Ser	Asp	Glu
312			50				55		60
314	gtg	gct	gaa	ttg	ccc	agg	gaa	gaa	gca
315	Val	Ala	Glu	Leu	Pro	Arg	Glu	Glu	Ala
316	65					70			75
318	aag	aat	ctt	aca	cca	tat	gtg	gct	att
319	Lys	Asn	Leu	Thr	Pro	Tyr	Val	Ala	Ile
320						85			90
322	aaa	gaa	cag	cag	ttt	agg	gag	tgg	ttt
323	Lys	Glu	Gln	Gln	Phe	Arg	Glu	Trp	Phe
324			100				105		110
326	aga	tgg	aag	att	cag	gag	tcc	ata	gaa
327	Arg	Trp	Lys	Ile	Gln	Glu	Ser	Ile	Glu
328			115				120		125
330	gag	att	gaa	aag	gtc	cac	aga	ggc	tgc
331	Glu	Ile	Glu	Lys	Val	His	Arg	Gly	Cys
332			130				135		140
334	ggc	tcc	act	ggc	atc	ctg	tct	gtc	att
335	Gly	Ser	Thr	Gly	Ile	Leu	Ser	Val	Ile
336	145					150			155
338	aca	gca	ggg	ctg	agc	ctg	agc	att	act
339	Thr	Ala	Gly	Leu	Ser	Leu	Ser	Ile	Thr
340						165			170
342	ata	gca	tct	gcc	acg	gct	ggg	atc	gcc
343	Ile	Ala	Ser	Ala	Thr	Ala	Gly	Ile	Ala
344			180				185		190
346	tac	aca	agg	tca	gca	gaa	ctc	aca	gcc
347	Tyr	Thr	Arg	Ser	Ala	Glu	Leu	Thr	Ala
348			195				200		205
350	act	gac	caa	ttg	gag	gca	tta	agg	gac
351	Thr	Asp	Gln	Leu	Glu	Ala	Leu	Arg	Asp
352			210				215		220
354	aat	gtg	ctt	tcc	ttt	gca	ctt	gat	ttt
355	Asn	Val	Leu	Ser	Phe	Ala	Leu	Asp	Phe
356	225					230			235
358	gcg	aat	gat	gtc	cat	aca	ctc	agg	aga
359	Ala	Asn	Asp	Val	His	Thr	Leu	Arg	Arg
360						245			250
362	cct	ttg	att	gct	tgg	cga	tat	gta	cct
363	Pro	Leu	Ile	Ala	Trp	Arg	Tyr	Val	Pro
364			260				265		270

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/835,996

DATE: 10/16/2001

TIME: 15:43:13

Input Set : A:\35915a.seq.txt

Output Set: N:\CRF3\10162001\I835996.raw

L:22 M:270 C: Current Application Number differs, Replaced Current Application No  
L:22 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:171 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:1372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:1373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:1385 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:1393 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:1572 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13  
L:1572 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13  
L:1580 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13  
L:1580 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13  
L:1592 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13  
L:1592 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13  
L:2747 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22  
L:3139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26  
L:6760 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:34  
L:6760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34  
L:6911 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:36  
L:6911 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36  
L:7554 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44

*Enumerate "n's"*